Published in final edited form as:
Am J Cardiol. 2012 December 15; 110(12): 1809–1813. doi:10.1016/j.amjcard.2012.08.015.

Reliability of Peak Exercise Testing in Patients with Heart Failure with Preserved Ejection Fraction

Jessica M. Scott, PhD[a], **Mark J. Haykowsky, PhD**[b], **Joel Eggebeen, MS**[c], **Timothy M. Morgan, PhD**[d], **Peter H. Brubaker, PhD**[e], and **Dalane W. Kitzman, MD**[c]

[a]NASA Johnson Space Center; Universities Space Research Association, Houston, Texas, USA

[b]University of Alberta, Edmonton, Alberta, Canada

[c]Department of Internal Medicine, Sections on Cardiology and Geriatrics, Wake Forest University School of Medicine, Winston-Salem, North Carolina, USA

[d]Division of Public Health Sciences, Wake Forest University Health Sciences, Winston-Salem, North Carolina, USA

[e]Department of Health and Exercise Science, Wake Forest University, Winston-Salem, North Carolina, USA

Abstract

Exercise intolerance is the primary symptom among heart failure patients with preserved ejection fraction (HFpEF), is a major determinant of their reduced quality of life, and an important outcome in clinical trials. Although cardiopulmonary exercise testing (CPET) provides peak and submaximal diagnostic indices, the reliability of peak treadmill CPET in patients ≥ 55 years of age with HFpEF has not been examined. Two CPETs were performed in 52 HFpEF patients (age 70 ± 7 years). The two tests were separated by an average of 23 ± 13 days (median: 22 days) and performed under identical conditions, with no intervention or change in status between visits except for initiation of a placebo run-in. A multi-step protocol for patient screening, education, and quality control was utilized. Mean peak VO_2 was similar on test 1 and test 2 (14.4 ± 2.4 vs. 14.3 ± 2.3 ml/kg/min). The correlation coefficients and intraclass correlations (ICC) from the testing days were as follows: VO_2 r = 0.85, p < 0.001, ICC = 0.855; ventilatory anaerobic threshold r= 0.79, p < 0.001, ICC= 0.790; VE/VCO_2 slope r = 0.87, p < 0.001, ICC = 0.864; HR r = 0.94, p < 0.001, ICC = 0.938. These results challenge conventional wisdom that serial baseline testing is required in clinical trials with exercise capacity outcomes. In conclusion, in women and men with HFpEF and severe physical dysfunction, key submaximal and peak exercise testing variables exhibited good reliability and were not significantly altered by a learning effect or placebo administration.

Correspondence: Dalane W. Kitzman, MD, Cardiology and Geriatrics Wake Forest University School of Medicine Winston-Salem, NC 27157; dkitzman@wfubmc.edu, phone: 336-716-3274, fax: 336-716-3274.

Conflicts of Interest: Dr. Kitzman is a consultant for Relypsa Inc., Boston Scientific Corp., and Abbot, has received grant support from Novartis, and owns stock in Gilead Sciences. Dr. Brubaker is a consultant to and has received research support from Boston Scientific Corp. All other authors report no relationships to disclose.

Keywords

heart failure preserved ejection fraction; cardiopulmonary exercise testing; test reproducibility

INTRODUCTION

Non-invasive assessment of exercise intolerance using cardiopulmonary exercise testing (CPET) provides crucial diagnostic data for the evaluation and treatment of patients with heart failure (HF), and is commonly used as a clinical endpoint in HF therapeutic trials.[1–3] CPET guidelines have been developed and regularly updated by expert working groups in order to standardize both testing procedures and interpretation of exercise data.[4,5] These guidelines commonly recommend that at least two CPETs be conducted: the first for habituation/learning purposes; the second for acquisition of baseline data.[5,6] The stated purpose of this multiple test approach is to reduce test variability. However, this adds significantly to participant burden, and study time and cost. Furthermore, although the reliability of CPET has been examined in symptomatic HF and reduced EF (HFrEF),[4] there are relatively few data verifying the reliability of peak exercise tests in patients ≥55 years, women, and patients with severely reduced exercise function who might benefit most by a habituation session. The latter patients are typical for the entity of HF with preserved EF (HFpEF), in which very few reliability data have been reported.[7] Importantly, the impact of a placebo run-in period on CPET variables in this population, relevant to pharmacological interventions, has not previously been investigated. The purpose of this study was to determine the reliability of a treadmill exercise test with a placebo run-in period for assessment of peak and submaximal indices of exercise tolerance in women and men ≥55 years with HFpEF utilizing participant screening, well-trained testing personnel, and adherence to stringent CPET criteria.

METHODS

The study protocol was approved by the Wake Forest University Medical Center Institutional Review Board, and written informed consent was obtained from all participants. This report is from the screening visit (test 1) and baseline visit (test 2) of a randomized, double-blind, placebo controlled pharmacological clinical trial whose primary outcome is exercise capacity. Participant eligibility was assessed at an initial screening visit, which included medical history, physical examination, blood labs, echocardiogram to document ejection fraction and evidence of diastolic dysfunction, and peak CPET (test 1). Approximately one week later, participants returned for a second screening visit. If the initial screening visit blood labs were within inclusion criteria, participants entered into a two week placebo run-in period. At the subsequent baseline visit, if compliance to the placebo run-in by pill count was at least 80%, participants completed all baseline study assessments, including a second peak CPET (test 2). Patients had NYHA class II–III of at least 1 month duration with preserved resting ejection fraction, echocardiographic evidence of diastolic dysfunction (early diastolic lateral mitral annular velocity ≥9 cm/sec.), and no evidence of significant anemia or coronary artery, valvular, infiltrative, pericardial, pulmonary, or renal disease. Additionally, only participants with evidence of severe exercise intolerance (exercise time of at least 2 minutes but not greater than 12 minutes for women or 14 minutes for men) were enrolled. Potential participants were initially screened for inclusion by retrospective review of clinic visits and hospital discharge records at the Wake Forest University Medical Center. Of the resultant 273 preliminarily eligible participants, 66 were scheduled for an initial screening visit, and 52 met the study criteria and were enrolled.

The multi-step screening, education, and quality control CPET protocol is outlined in Appendix 1. Briefly, exercise testing was performed on a treadmill using the modified Naughton protocol and conducted by the same trained, experienced individual with master's degree in exercise physiology. Prior to each test, participants were screened for any recent acute illness or injury or change in status and were given detailed, standardized, verbal instructions. During the test, participants were encouraged to walk for as long as possible and, once fatigued, to attempt the subsequent stage, resulting in a peak, symptom-limited exhaustive effort. Complete calibration of gas concentrations and flow were performed before each test. Metabolic gas exchange was measured continuously during exercise and averaged over 15-second intervals (Medgraphics Ultima, Medical Graphics Corp., St. Paul, Minnesota). Peak VO_2 was defined as the average of the 2 highest VO_2 values for a given 15-second interval within the last 90 seconds of exercise. Ventilatory anaerobic threshold (VAT) was calculated using standard methods.[8] The VE/VCO_2 slope was determined by measuring the slope across the entire course of exercise.[9] Oxygen uptake efficiency slope (OUES) was determined measuring the slope of VO_2 (ml/min) and $\log_{10}VE$ (L/min) across the entire course of exercise.[10] All tests were evaluated by the exercise physiologist for validity based on a combination of symptoms (dyspnea on exertion and/or fatigue as primary endpoints) and physiological measures (respiratory exchange ratio, heart rate, minute ventilation, respiratory rate) and confirmed if necessary by a cardiologist. To avoid bias, the tester intentionally did not view the results of test 1 immediately before conducting test 2. Additionally, validation of the gas exchange unit was conducted prior to each testing day by performing a 2-stage, incremental exercise test in a known, constant test subject. In order to be validated, VO_2 and VCO_2 needed to be within 2 standard deviations of the usual, known value at each workload.

Means are expressed using either ± standard deviation or with 10th and 90th percentiles. Within-subject variability from test 1 to test 2 was quantified by the within-subject absolute change (i.e., either increase or decrease from test 1 to test 2). A paired t-test was used to test for a statistical difference between the test 1 and test 2 values. Bland-Altman plot was used to visually assess whether the magnitude of within-subject variability in peak VO_2 varied with the magnitude of the measurements.[11] Reliability of the CPET outcome measures between the screening visit and baseline visit was tested by Pearson correlations and intraclass correlations. One-sided 95% confidence intervals of Pearson correlations were made using Fisher's z transformation. For each variable, the coefficient of variation (CV) was defined as the (within-subject standard deviation/within-subject mean) × 100%. Significance was set at $p<0.05$.

RESULTS

Participants were predominantly white women with a history of hypertension and NYHA functional class II symptoms (Table 1). Test 1 and test 2 were separated by an average of 23 ± 13 days. Compliance by pill count to the placebo run-in period was excellent (96%). No adverse events occurred during or after any peak CPET. In three participants, the screening visit CPET was not considered to be a valid peak test and was repeated two weeks later. The reasons were: 1) abnormal gait and coordination resulting in early fatigue and test termination in two participants, and 2) test ended primarily due to knee pain in one participant. The two participants with abnormal gait and coordination were coached on proper mechanics immediately following the screening visit CPET with good results. The participant with initial knee pain returned following a bout of NSAIDS and completed the repeated CPET with no pain. In each case, the repeated screening visit CPET was considered to be valid and was used in this analysis.

Peak VO_2 and VAT were nearly identical on the 2 tests. Within-subject peak VO_2 demonstrated high agreement between tests as shown by Bland-Altman analysis (Figure 1). Reliability of CPET parameters was excellent (Table 2), and was similar across age and gender (Tables 3 and 4). Exercise time increased significantly from test 1 to test 2. Likewise, VE/VCO_2 slope was significantly greater in test 2 than test 1 (Table 5).

DISCUSSION

Few data are available regarding the reliability of CPET in women, and specifically in patients ≥55 years of age with HFpEF despite the fact that HFpEF occurs in approximately 50% or more of HF patients in the community, and the proportion is higher among women and individuals >55 years.[12–14] Since HFpEF patients are predominantly ≥55 years and female, they may be expected to be less familiar with exercise testing, to have greater variability, and to have a greater need for "learning" or practice sessions. Moreover, the impact of a placebo run-in period on CPET variables in this population, relevant to pharmacological interventions, has not previously been investigated. Accordingly, in this study we investigated the reliability of CPET-derived variables specifically in 52 patients ≥55 years with HFpEF who underwent 2 identical CPETs approximately 3 weeks apart with a placebo run-in period between tests. We utilized a multi-step screening, education, and quality control protocol to optimize reliability. The major new findings of this study were 1) No change in the mean value or significant within-subject variability in peak VO_2 or VAT between tests, suggesting no significant learning or placebo effect, 2) Good correlations in all CPET parameters between tests, with no difference across age and gender. Thus, in patients with HFpEF ≥55 years of age, there may be no need to perform more than one CPET to obtain valid, baseline peak VO_2 and other CPET results.

Numerous studies have shown that peak VO_2 is a key outcome for assessing the prognosis of patients with HF, and is commonly used as a clinical endpoint in HF therapeutic trials, including those with HFpEF.[1,2,15] Current cardiology guidelines recommend that at least two CPETs be conducted based on prior work demonstrating that peak VO_2 can vary significantly over serial tests.[5] For example, Elborn et al.[16] performed three consecutive treadmill tests separated by two weeks on 30 subjects with HFrEF. The mean peak VO_2 improved significantly from the first to the second test (14.1 vs. 14.9 mL/kg/min, $p<0.005$), while the average within-subject CV for the three tests was 6%. The authors concluded that a single baseline test was insufficient and suggested that the performance of at least two tests be conducted for clinical research applications. In contrast, Bensimhon et al.[4] performed two baseline peak treadmill tests on 398 men and women with symptomatic HFrEF and found that there was no significant change in peak VO_2 (15.2 ± 5.0 vs. 15.2 ± 5.0 mL/kg/min; $p=0.78$). Although within-subject variability of peak VO_2 was significant, the authors concluded that in multi-center trials, there is no need to perform more than 1 baseline CPET. In the only study to examine HFpEF to date, Marburger et al.[7] examined the reproducibility of a cycle ergometer CPET in 5 HFpEF patients and 4 HFrEF patients. The authors concluded that the reliability of peak testing was good in both HFpEF and HFrEF, and that only one test was necessary. However, the small number of patients with HFpEF significantly limits the widespread applicability of results. In the present study with 52 HFpEF patients, for peak VO_2, there was good correlation between tests ($r = 0.85$; ICC = 0.855), the overall group means on test 1 and test 2 were essentially identical, the percentage of subjects who decreased or increased on test 2 were similar, Bland-Altman plots demonstrated high agreement between the 2 tests, and the within-subject variability was low (average CV <5%). Taken together, these findings further suggest that a single baseline test is sufficient to measure peak VO_2 for the evaluation of therapy or assessment of prognosis.

A large body of research has also evolved regarding CPET variables other than peak VO_2. Responses such as VAT, VE/VCO_2 slope, OUES, and exercise time have been used with greater frequency to classify functional limitations and to stratify risk in patients with HF.[5] VAT and OUES are independent of effort and protocol and, similar to peak VO_2, have been shown to be highly reproducible.[4,10] In this study, we demonstrate that VAT and OUES are also reproducible with repeat testing as OUES and VAT had good correlation between tests and nearly identical mean values on test 1 and 2. As seen previously in HFrEF,[4,16,17] we found that the mean exercise time increased from test 1 to test 2, with 54% of subjects increasing on test 2. Additionally, mean values for VE/VCO_2 slope increased from test 1 to test 2, and there were a greater number of subjects who increased on test 2 than decreased. The reasons for discordant findings for VE/VCO_2 slope are uncertain; however, may be related to daily variations in ventilatory oscillation. Guazzi et al.[18] reported that the prevalence of abnormal ventilatory oscillation in patients with HFpEF approximates 30%, which could affect VE/CO_2.

To achieve these results and be able to conduct a single, valid baseline CPET in longitudinal studies, there are several important aspects to consider (see appendix). These include: standardized participant instructions, screening for intercurrent acute illness or injury or change in condition, frequent calibration and validation of the gas exchange unit, strong encouragement to reach a peak level, and evaluation of reasons for termination and physiological responses. CPET that do not meet these criteria will likely have reduced reliability. Although inter-observer variability was not assessed in the present study given that each test was conducted by the same experienced individual, high reproducibility is obtainable with appropriate training of laboratory personnel.[4,7] These results challenge conventional wisdom that serial baseline testing is required in clinical trials with exercise capacity outcomes, and suggest that any incremental gain be balanced with the increase in participant burden and cost.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Funding: This study was supported in part by N.I.H. grants R37AG18915 and P30AG21332 and by an investigator-initiated grant from Novartis Pharmaceuticals.

References

1. Haykowsky MJ, Brubaker PH, John JM, Stewart KP, Morgan TM, Kitzman DW. Determinants of exercise intolerance in elderly heart failure patients with preserved ejection fraction. J Am Coll Cardiol. 2012; 58:265–274. [PubMed: 21737017]
2. Kitzman DW, Daniel KR. Diastolic heart failure in the elderly. Clin Geriatr Med. 2007; 23:83–106. [PubMed: 17126756]
3. Haykowsky M, Brubaker P, Kitzman D. Role of physical training in heart failure with preserved ejection fraction. Curr Heart Fail Rep. 2012; 9:101–106. [PubMed: 22430146]
4. Bensimhon DR, Leifer ES, Ellis SJ, Fleg JL, Keteyian SJ, Pina IL, Kitzman DW, McKelvie RS, Kraus WE, Forman DE, Kao AJ, Whellan DJ, O'Connor CM, Russell SD. Reproducibility of peak oxygen uptake and other cardiopulmonary exercise testing parameters in patients with heart failure (from the Heart Failure and A Controlled Trial Investigating Outcomes of exercise traiNing). Am J Cardiol. 2008; 102:712–717. [PubMed: 18773994]
5. Arena R, Myers J, Williams MA, Gulati M, Kligfield P, Balady GJ, Collins E, Fletcher G. Assessment of functional capacity in clinical and research settings: a scientific statement from the American Heart Association Committee on Exercise, Rehabilitation, and Prevention of the Council

on Clinical Cardiology and the Council on Cardiovascular Nursing. Circulation. 2007; 116:329–343. [PubMed: 17576872]

6. Myers J, Froelicher VF. Optimizing the exercise test for pharmacological investigations. Circulation. 1990; 82:1839–1846. [PubMed: 2225380]
7. Marburger CT, Brubaker PH, Pollock WE, Morgan TM, Kitzman DW. Reproducibility of cardiopulmonary exercise testing in elderly patients with congestive heart failure. Am J Cardiol. 1998; 82:905–909. [PubMed: 9781977]
8. Beaver WL, Wasserman K, Whipp BJ. A new method for detecting anaerobic threshold by gas exchange. J Appl Physiol. 1986; 60:2020–2027. [PubMed: 3087938]
9. Arena R, Myers J, Aslam SS, Varughese EB, Peberdy MA. Technical considerations related to the minute ventilation/carbon dioxide output slope in patients with heart failure. Chest. 2003; 124:720–727. [PubMed: 12907564]
10. Baba R, Nagashima M, Goto M, Nagano Y, Yokota M, Tauchi N, Nishibata K. Oxygen uptake efficiency slope: a new index of cardiorespiratory functional reserve derived from the relation between oxygen uptake and minute ventilation during incremental exercise. J Am Coll Cardiol. 1996; 28:1567–1572. [PubMed: 8917273]
11. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. Lancet. 1986; 1:307–310. [PubMed: 2868172]
12. Kitzman DW, Gardin JM, Gottdiener JS, Arnold A, Boineau R, Aurigemma G, Marino EK, Lyles M, Cushman M, Enright PL. Importance of heart failure with preserved systolic function in patients > or = 65 years of age. CHS Research Group. Cardiovascular Health Study. Am J Cardiol. 2001; 87:413–419. [PubMed: 11179524]
13. Vasan RS, Larson MG, Benjamin EJ, Evans JC, Reiss CK, Levy D. Congestive heart failure in subjects with normal versus reduced left ventricular ejection fraction: prevalence and mortality in a population-based cohort. J Am Coll Cardiol. 1999; 33:1948–1955. [PubMed: 10362198]
14. Bhatia RS, Tu JV, Lee DS, Austin PC, Fang J, Haouzi A, Gong Y, Liu PP. Outcome of heart failure with preserved ejection fraction in a population-based study. N Engl J Med. 2006; 355:260–269. [PubMed: 16855266]
15. Haykowsky M, Brubaker P, Kitzman D. Role of Physical Training in Heart Failure with Preserved Ejection Fraction. Curr Heart Fail Rep. 2011:2011.
16. Elborn JS, Stanford CF, Nicholls DP. Reproducibility of cardiopulmonary parameters during exercise in patients with chronic cardiac failure. The need for a preliminary test. Eur Heart J. 1990; 11:75–81. [PubMed: 2106439]
17. Sullivan M, Genter F, Savvides M, Roberts M, Myers J, Froelicher V. The reproducibility of hemodynamic, electrocardiographic, and gas exchange data during treadmill exercise in patients with stable angina pectoris. Chest. 1984; 86:375–382. [PubMed: 6467998]
18. Guazzi M, Myers J, Peberdy MA, Bensimhon D, Chase P, Arena R. Exercise oscillatory breathing in diastolic heart failure: prevalence and prognostic insights. Eur Heart J. 2008; 29:2751–2759. [PubMed: 18836201]

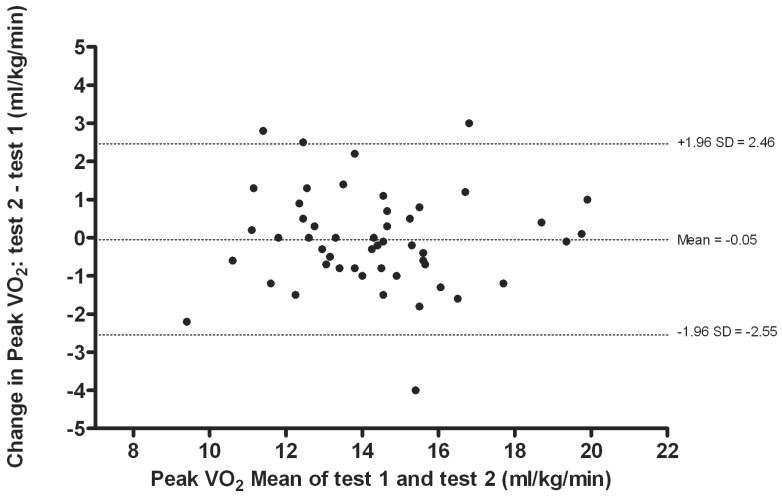

Figure 1.
Bland-Altman plot for change in peak VO_2 versus mean peak VO_2 for test 1 and test 2.

Table 1

Participant characteristics (n=52)

Age (years)	70 ± 7
Men	10 (19%)
Women	42 (81%)
White	31 (60%)
Black	21 (40%)
Weight (kg)	89 ± 17
Systolic blood pressure (mm Hg)	129 ± 18
Diastolic blood pressure (mm Hg)	74 ± 10
Body mass index (kg/m)	33.5 ± 5.7
New York Heart Association functional class	
II	37 (71%)
III	15 (29%)
Doppler diastolic function	
Abnormal relaxation	38 (73%)
Pseudonormal	11 (21%)
Undetermined (atrial fibrillation)	3 (6%)
Diabetes Mellitus	21 (40%)
Hypertension (by history)	50 (96%)
Medications	
angiotensin-converting enzyme inhibitors	20 (38%)
Digoxin	1 (2%)
Diuretics	37 (71%)
Beta-blockers	24 (46%)
Calcium-channel blockers	19 (37%)
Peak oxygen consumption, test 1 (ml/kg/min)	14.4 ± 2.4

Values are expressed as n (%) or mean ± SD.

Table 2

Reliability of cardiopulmonary exercise testing

	Test 1	Test 2	R	1-sided 95% CI	ICC	P-Value
Peak Exercise (treadmill)						
Time (seconds)	577 ± 111	598 ± 123	0.914	0.866	0.895	<0.001
Indexed peak oxygen uptake (ml/kg/min)	14.4 ± 2.4	14.3 ± 2.3	0.853	0.775	0.855	<0.001
Absolute peak oxygen uptake (ml/min)	1266 ± 300	1260 ± 290	0.918	0.872	0.919	<0.001
Heart rate (beats/min)	128 ± 25	127 ± 24	0.938	0.903	0.938	<0.001
Systolic blood pressure (mmHg)	170 ± 18	167 ± 19	0.772	0.659	0.763	<0.001
Diastolic blood pressure (mmHg)	74 ± 10	72 ± 8	0.746	0.622	0.726	<0.001
Peak carbon dioxide production (ml/min)	1427 ± 346	1431 ± 346	0.893	0.834	0.895	<0.001
Respiratory exchange ratio	1.13 ± 0.06	1.14 ± 0.07	0.623	0.458	0.622	<0.001
Ventilation per carbon dioxide slope	30.5 ± 4.5	31.1 ± 4.5	0.872	0.802	0.864	<0.001
Oxygen uptake at ventilatory threshold (ml/min)	845 ± 213	840 ± 212	0.787	0.678	0.790	<0.001
Oxygen uptake efficiency slope (ml/min*L/min)	1567 ± 449	1553 ± 450	0.869	0.798	0.871	<0.001

P-Value corresponds to the R value.

Table 3

Effects of age on cardiopulmonary exercise reliability

	Age < 70 (n=28)				Age ≥ 70 (n=24)			
	Test 1	**Test 2**	**R**	**ICC**	**Test 1**	**Test 2**	**R**	**ICC**
Exercise time (sec)	593 ± 99	617 ± 100	0.90	0.878	557 ± 124	577 ± 145	0.92	0.904
Oxygen uptake (ml/kg/min)	14.2 ± 2.5	14.2 ± 2.5	0.89	0.894	14.6 ± 2.3	14.5 ± 2.3	0.80	0.806
Oxygen uptake (ml/min)	1305 ± 262	1295 ± 228	0.91	0.900	1220 ± 339	1220 ± 349	0.93	0.931
Heart rate (beats/min)	132 ± 25	132 ± 22	0.93	0.920	124 ± 26	121 ± 26	0.96	0.954
Systolic blood pressure (mmHg)	172 ± 17	169 ± 17	0.80	0.798	167 ± 19	164 ± 22	0.74	0.734
Diastolic blood pressure (mmHg)	76 ± 9	74 ± 8	0.77	0.742	72 ± 11	70 ± 8	0.71	0.697
Carbon dioxide production (ml/min)	1448 ± 286	1462 ± 257	0.87	0.867	1401 ± 410	1395 ± 430	0.91	0.910
Respiratory exchange ratio	1.12 ± 0.06	1.13 ± 0.06	0.68	0.661	1.15 ± 0.06	1.14 ± 0.08	0.60	0.578
Ventilation per carbon dioxide slope	29.7 ± 4.7	30.5 ± 4.9	0.88	0.875	31.3 ± 4.2	31.8 ± 4.0	0.85	0.843
Oxygen uptake at ventilatory threshold (ml/min)	861 ± 190	860 ± 163	0.76	0.755	827 ± 239	818 ± 258	0.81	0.814
Oxygen uptake efficiency slope	1642 ± 340	1641 ± 403	0.77	0.762	1480 ± 545	1451 ± 487	0.94	0.932

Table 4

Effects of gender on cardiopulmonary exercise reliability

	Women (n=42)				Men (N=10)			
	Test 1	**Test 2**	**R**	**ICC**	**Test 1**	**Test 2**	**R**	**ICC**
Exercise time (sec)	568 ± 107	590 ± 116	0.92	0.902	611 ± 126	632 ± 151	0.89	0.877
Oxygen uptake (ml/kg/min)	13.8 ± 1.9	13.9 ± 1.8	0.79	0.795	16.8 ± 2.7	16.3 ± 3.2	0.88	0.865
Oxygen uptake (ml/min)	1183 ± 234	1187 ± 221	0.89	0.888	1614 ± 306	1571 ± 348	0.89	0.885
Heart rate (beats/min)	133 ± 24	132 ± 21	0.93	0.927	110 ± 24	107 ± 26	0.94	0.934
Systolic blood pressure (mmHg)	171 ± 18	168 ± 20	0.80	0.789	165 ± 18	161 ± 16	0.61	0.616
Diastolic blood pressure (mmHg)	74 ± 9	72 ± 8	0.71	0.694	74 ± 12	72 ± 9	0.86	0.840
Carbon dioxide production (ml/min)	1321 ± 251	1338 ± 256	0.84	0.842	1870 ± 348	1824 ± 406	0.84	0.840
Respiratory exchange ratio	1.12 ± 0.06	1.13 ± 0.07	0.55	0.550	1.16 ± 0.06	1.17 ± 0.07	0.84	0.852
Ventilation per carbon dioxide slope	30.7 ± 4.5	31.5 ± 4.6	0.86	0.852	29.4 ± 4.9	29.7 ± 4.4	0.91	0.908
Oxygen uptake at ventilatory threshold (ml/min)	790 ± 170	783 ± 157	0.71	0.716	1070 ± 231	1074 ± 252	0.68	0.705
Oxygen uptake efficiency slope	1455 ± 333	1460 ± 378	0.81	0.811	2040 ± 577	1944 ± 532	0.89	0.886

Table 5

Test-retest variability of peak exercise variables

	Test 1	Test 2	P-Value	Absolute Change (2 – 1)	% 2 > 1	% 1 > 2	cv
Exercise Time (seconds)	577 (411, 720)	598 (439, 720)	0.003*	22 (−50, 90)	54	25	4.9%
Indexed oxygen uptake (ml/kg/min)	14.4 (11.1, 18.0)	14.3 (11.6, 18.0)	0.79	−0.05 (−1.50, 1.37)	40	52	4.9%
Ventilatory threshold (ml/min)	845 (581, 1177)	840 (603, 1064)	0.81	−5 (−175, 190)	45	53	9.3%
Ventilation per carbon dioxide slope	30.5 (25.3, 36.3)	31.1 (24.9, 36.8)	0.04*	0.7 (−2.2, 3.8)	65	35	4.5%
Respiratory exchange ratio	1.13 (1.07, 1.21)	1.14 (1.06, 1.24)	0.55	0.01 (−0.07, 0.07)	50	44	3.0%
Oxygen uptake efficiency slope (ml/min*L/min)	1567 (1074, 2071)	1553 (1107, 2076)	0.67	−14 (−243, 341)	40	60	8.5%

Values expressed as mean (10^{th} percentile, 90^{th} percentile) unless noted otherwise. P-Value corresponds to a paired t-test of the null hypothesis that the mean of test 2 minus the mean of test 1 = 0.

*paired t-test of the null hypothesis that the mean of test 2 minus the mean of test 1 =0.